Journal of Peptide Science
J. Pept. Sci. 2007; **13**: 314–326
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.843

JOURNAL OF PEPTIDE SCIENCE

The sequence TGAAKAVALVL from glyceraldehyde-3-phosphate dehydrogenase displays structural ambivalence and interconverts between α-helical and β-hairpin conformations mediated by collapsed conformational states

SUNITA PATEL,[a] PETETY V. BALAJI[b] and YELLAMRAJU U. SASIDHAR[a,b]*

[a] Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India
[b] School of Biosciences & Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

Received 9 January 2007; Accepted 13 January 2007

Abstract: The peptide TGAAKAVALVL from glyceraldehyde-3-phosphate dehydrogenase adopts a helical conformation in the crystal structure and is a site for two hydrated helical segments, which are thought to be helical folding intermediates. Overlapping sequences of four to five residues from the peptide, sample both helical and strand conformations in known protein structures, which are dissimilar to glyceraldehyde-3-phosphate dehydrogenase suggesting that the peptide may have a structural ambivalence. Molecular dynamics simulations of the peptide sequence performed for a total simulation time of 1.2 μs, starting from the various initial conformations using GROMOS96 force field under NVT conditions, show that the peptide samples a large number of conformational forms with transitions from α-helix to β-hairpin and vice versa. The peptide, therefore, displays a structural ambivalence. The mechanism from α-helix to β-hairpin transition and vice versa reveals that the compact bends and turns conformational forms mediate such conformational transitions. These compact structures including helices and hairpins have similar hydrophobic radius of gyration (R_{gh}) values suggesting that similar hydrophobic interactions govern these conformational forms. The distribution of conformational energies is Gaussian with helix sampling lowest energy followed by the hairpins and coil. The lowest potential energy of the full helix may enable the peptide to take up helical conformation in the crystal structure of the glyceraldehyde-3-phosphate dehydrogenase, even though the peptide has a preference for hairpin too. The relevance of folding and unfolding events observed in our simulations to hydrophobic collapse model of protein folding are discussed. Copyright © 2007 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: helix to hairpin transition; bend; turn; peptide model; structural ambivalence; GROMACS; protein folding; molecular dynamics

INTRODUCTION

Recently, it has been observed that identical short sequence fragments from different proteins adopt different secondary structures [1–6]. For instance, Mezei [4] found several instances of identical hexamers and heptamers, which sample α-helical and β-sheet conformations in different proteins. A sequence, which adopts α-helical conformation in one protein and β-strand conformation in another protein is termed as a chameleon sequence [4,7]. In general, if a sequence adopts different conformations in different proteins, such a sequence is termed as a structurally ambivalent peptide (SAP) [5,6]. Recently, Kuznetsov and Rackovsky [6] performed a thorough statistical analysis of SAPs in protein structures. They find that SAPs have simultaneous intrinsic preference for two distinct types of backbone conformations. They particularly found that sequence fragments, that adopt a helical conformation in one protein and sheet conformation in another protein, have the lowest sequence complexity and such sequences are rich in L, V and A residues, a conclusion shared by Mezei [4] and Zhou *et al.* [5]. Kuznetsov and Rackovsky [6] further found that the choice of secondary structure sampled by SAPs depends on the flanking residues of the sequence and also on the global context of the protein. In other words, while SAPs may have equal intrinsic conformational preference for different secondary structures, the choice of the secondary structure actually adopted is determined by the protein context.

Recently, Young *et al.* [8] developed a program called an ambivalent structural predictor to identify structurally ambivalent sequence elements. They analyzed the sequence of 16 proteins that are known to undergo conformational switching as a manifestation of their biological activity. They found a correlation between structurally ambivalent regions and conformational switches. Further, structural ambivalence may have implications for the development of amyloid related diseases [9–12]. Alpha to beta conformational switching

* Correspondence to: Yellamraju U. Sasidhar, Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India; e-mail: sasidhar@iitb.ac.in

from nonnative α-helix to β-sheet conformation is known to occur during the refolding of β-lactoglobulin, which is a predominantly β-sheet protein [13–15]. During the refolding of lysozyme, an overshoot of helical CD is observed during the early folding and subsequently native helicity is obtained [16]. Also in this case, an α to β conformational change is involved. Thus, there is a need to understand structural ambivalence in view of its implications for conformational switching, conformational diseases and protein folding.

In this paper, we present an example of a sequence that displays structural ambivalence between α-helical and β-hairpin conformations in molecular dynamics simulations performed for a total simulation time of 1.2 μs. The sequence TGAAKAVALVL from glyceraldehyde-3-phosphate dehydrogenase adopts helical conformation in the crystal structure of the protein. Further, the sequence contains hydrated helical segments, which are thought to be helical folding intermediates [17]. However, a search of the protein data bank reveals that overlapping peptide fragments of four and five residues from this sequence display structural ambiguity. That is, they sample both α-helical and β-strand conformational forms (Table 1). Hence, the sequence can be considered as a candidate for displaying α- to β- conformational transition. Since the sequence is shorter, it is amenable to molecular dynamics simulation studies with significant lengths of simulation times. We consider three starting conformations: helical conformation as observed in the crystal structure of the glyceraldehyde-3-phosphate dehydrogenase, polyproline II conformation and fully extended conformation. We considered polyproline II conformation as a model for the unfolded state, since some of the recent spectroscopic studies of unfolded polyalanine peptides suggest that they sample polyproline II helical conformation [18–21]. Recent simulations also find that unfolded state of polyalanine is a segmented polyproline II helical conformation, which is also a preferred conformation [22,23]. In our molecular dynamics simulations, starting from the helical, extended and polyproline II conformational forms, the sequence samples both α-helical and β-hairpin conformational forms and we find that the compact turns and bends conformational forms mediate these conformational changes. Thus, the sequence displays structural ambivalence and the possible reasons for it are discussed. Further, the relevance of folding and unfolding events observed in our simulations to hydrophobic collapse model are also discussed.

Table 1 Structural ambiguity displayed by overlapping tetra and penta peptide fragments from the sequence TGAAKAVALVL from glyceraldehyde-3-phosphate dehydrogenase as observed in different protein structures

Peptide fragment	PDB id	Start residue	Secondary structure[a]	Sequence identity (%)[b]
TGAAK	1K3T:B	226	CSHHH	51
	1P33:A	13	TTTTT	18
GAAKA	1DSS:G	208	THHHH	53
	1BKG:A	231	ESTTT	7
AAKAV	1AZY	413	HHHHH	14
AKAVA	1EXC:A	54	HHHHH	12
KAVAL	1N61:C	245	HHHHH	17
	1OOT:A	5	EEEES	4
AVALV	1SUF:A	532	HHHHH	10
	1JNF	95	EEEEE	4
VALVL	1JN0	215	HHHHC	59
	1UUS:A	362	EEEEE	8
TGAA	1R4W	171	HHHH	10
	1RM6:A	476	ECCC	9
GAAK	1K89	244	HHHH	15
	1RP1	416	EEEE	6
AAKA	1EDO	169	HHHH	16
	1XPK	84	SSSC	15
AKAV	1JDS	227	HHHH	8
	1UKK	6	EEEE	8
KAVA	1DTZ	147	HHHH	11
	1UZB	510	EEEE	6
AVAL	1YQG:A	203	HHHH	15
	1PPL:E	154	EEEC	10
VALV	1AXN	86	HHHH	14
	1Z2R	419	EEEE	16
ALVL	1J1W	626	HHHH	6
	1YQG:A	82	CEEE	15

[a] The secondary structure assignments are H, α-helix; E, extended β-strand; S, bend; T, turn; C, coil state.
[b] Sequence identity (%) is calculated using EMBOSS (http://www.ebi.ac.uk/emboss/align/) after the global alignment of glyceraldehyde-3-phosphate dehydrogenase with other protein sequences.

METHODS

MD simulations are performed on the peptide ^{208}TGAAKAVALVL^{218} from glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. The crystal structure of the protein is available (PDB code 1GD1). The ends of the peptides are protected by the CH_3CO- group (Ac) at the *N*-terminus and the —$NHCH_3$ group (NMe) at the *C*-terminus. In all the simulations reported, the peptide is given an initial conformation, which is either fully extended or a polyproline II or a native helical conformation. In fully extended conformation, all ϕ, ψ and side chain dihedral angles are set to 180° except for χ^1, which is set at 60°. The peptide in polyproline II conformation is generated using Deep View [24] by setting all backbone dihedral angles of the peptide from crystalline glyceraldehyde-3-phosphate dehydrogenase (PDB code 1GD1) to polyproline II values ($\phi = -76°$ and $\psi = 149°$).

The following abbreviations are used to indicate the simulations performed:

AH, simulation with initial helical conformation, which is its native conformation in the crystalline structure of glyceraldehyde-3-phosphate dehydrogenase.

DOI: 10.1002/psc

PP, simulation with initial polyproline II conformation of the peptide
ET, simulation with initial extended conformation of the peptide

One simulation run is performed for each initial conformation with a total simulation time of 1.2 μs. All the simulations are performed in cubic boxes with periodic boundary conditions. The cubic box sides in AH, PP and ET simulations are 4.24, 5.71 and 6.34 nm, respectively, and contained 2484, 5766 and 8328 SPC [25] water molecules, respectively, solvating the peptide.

The electrostatic interactions are treated by PME method [26,27] with a Coulomb cutoff of 1 nm, Fourier spacing of 0.12 nm and an interpolation order of 4. The van der Waals interactions are treated using Lennard-Jones potential and a switching function with a cutoff distance of 1 nm and a switching distance of 0.9 nm. The total charge of the system is +1 in electronic charge units. In all the simulations, a negative counter ion, Cl^- was added by replacing a water molecule to achieve electrical neutrality.

The potential energy of the peptide in water is minimized using the steepest descent algorithm with a tolerance of 100 kJ mol^{-1} nm^{-1} and convergence is obtained in all the cases. Subsequent to energy minimization, position restrained molecular dynamics is carried out for 50 ps using a reference temperature of 300 K. In this procedure, the atomic positions of the peptide are restrained and the water molecules are allowed to equilibrate around the peptide, to remove the solvent holes. Initial velocities required to start the procedure are generated conforming to Maxwell velocity distribution at 300 K. Following these equilibration procedures, MD is initiated. A time step of 2 fs is used for integrating the equations of motion. LINCS algorithm is used to constrain the bonds [28]. Coordinates are saved every 250 steps or 0.5 ps and velocities are saved every 500 steps or 1 ps. The peptide and the solvent are separately coupled to a Berendsen temperature bath [29] at 300 K using a time constant of 0.1 ps. All the simulations are done under NVT conditions.

MD simulations are performed using the software GROMACS (version 3.1.4) [30] and GROMOS96 (ffG43a1) force field [31,32] on dual Xeon processor-based machines running RedHat Linux 8 (http://www.redhat.com). Except for polar hydrogens like —NH, —OH and aromatic ring hydrogens, united atom approximation is used. The analysis tools provided by GROMACS software are used to analyze the data. The radius of gyration (R_g) refers to the R_g of the backbone atoms. The hydrophobic radius of gyration (R_{gh}) refers to the R_g of hydrophobic side chains excluding C_α atoms. For the secondary structure assignment, DSSP program [33] is used, which is implemented in GROMACS. According to DSSP, a coil conformation is defined by a low curvature region without any hydrogen bonds and a bend conformation at ith residue is defined involving i − 2, i − 1, i, i + 1 and i + 2 residues where the backbone changes direction by more than 70° at ith residue [33]. A bend may, in principle, develop into two consecutive turns with (i-1, i) and (i, i+1) as the corner residues. In a bridge conformation, a pair of hydrogen bonds form adjacent to each other involving a main-chain NH and CO groups connecting opposite strands [33]. A conformation of the peptide having bend(s) and turn(s) is referred to as bends and turns conformation.

Matlab (http://www.mathworks.com) and Octave (http://www.octave.org) are also used for some of the analyses. Most of the graphs and figures are generated using Xmgrace (http://plasma-gate.weizmann.ac.il/Grace), Matlab and VMD [34].

RESULTS

The Peptide Samples Compact and Extended Conformational Forms During the Simulation

Figure 1(a) shows the R_g of the peptide *versus* time. It is clear that R_g fluctuates between ∼1 and ∼0.5 nm. We note that α-helical and β-hairpin conformational forms have R_g values around ∼0.6 nm. Compact conformational forms like bends with bridge and turns with bridge have R_g values around ∼0.5 nm. The random coil conformation has an R_g value of about 1 nm and R_{gh} value of about 0.8 nm. From the Figure 1(a), it is, therefore, obvious that the R_g value fluctuates between values corresponding to compact conformational forms and extended random coil conformational forms. Thus, the graph shows the transitions between compact and extended conformational forms. A similar behavior is seen in the R_{gh} *versus* time graph (Figure 1(b)) suggesting that hydrophobic interactions may underlie these conformational transitions. In all the simulations, mean R_g and R_{gh} values are similar (Figure 1) suggesting that similar compact conformations are sampled in all the simulations.

The Peptide Shows Transitions Between α-Helical and Sheet Conformations

The number of residues in α-helical or sheet conformations are shown in Figure 2 as a function of time. As can be noticed from the plot, these conformations alternate in time. For example, in AH simulation, up to about ∼30 ns helical conformation is observed. Subsequently, β-sheet conformation is observed at about ∼40 ns. After about ∼53 ns, helical conformation is again observed. This kind of transitions can be seen to occur in all the simulations. Conformational transitions in more detail are shown in Figure 3.

The Peptide Samples Both α-Helical and β-Hairpin Conformations

Figure 3 shows the evolution of secondary structure as a function of time in a DSSP plot [33]. The residue numbers are indicated along the *y*-axis and the conformations are indicated by various colors. In AH simulation (Figure 3), the simulation starts with an all helix conformation. By about 20 ns, the peptide begins to start losing its helical structure from the *C*-terminal end. By about 30 ns, the helical structure is completely lost and a family of bends

DOI: 10.1002/psc

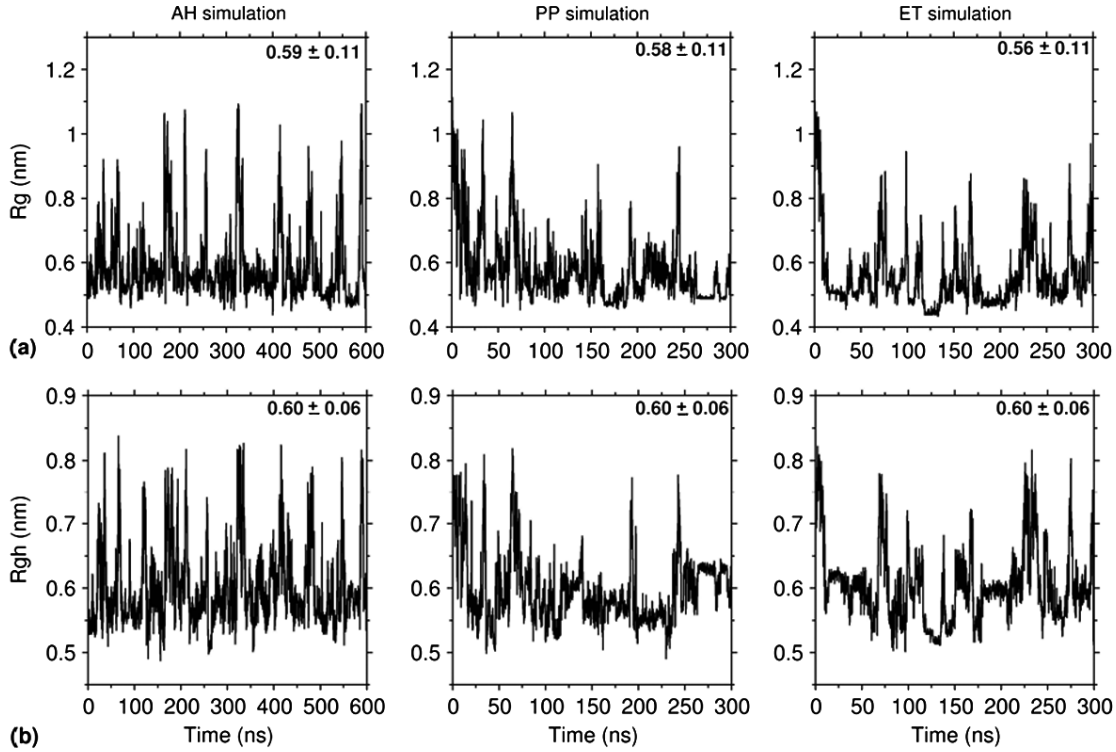

Figure 1 Radius of gyration (R_g) and hydrophobic radius of gyration (R_{gh}) of the peptide, as a function of time in AH simulation, PP simulation and ET simulation in panel (a) and panel (b), respectively. For visual clarity, running averages over 400 data points (corresponding to 200 ps) are plotted. The inset in both the panels shows the mean R_g and R_{gh} values along with standard deviations in nanometers. R_g and R_{gh} are defined in the section 'Methods'.

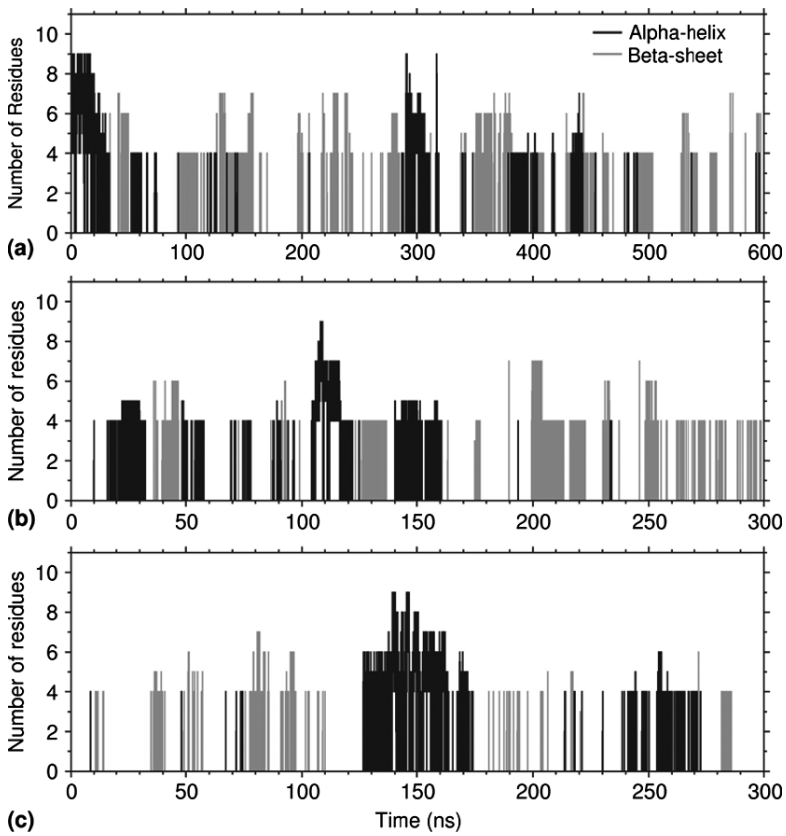

Figure 2 Number of residues in α-helix or β-sheet as a function of time in AH simulation (panel a), PP simulation (panel b) and ET simulation (panel c). The number of residues in α-helix or β-sheet conformation are computed by the program DSSP[33].

DOI: 10.1002/psc

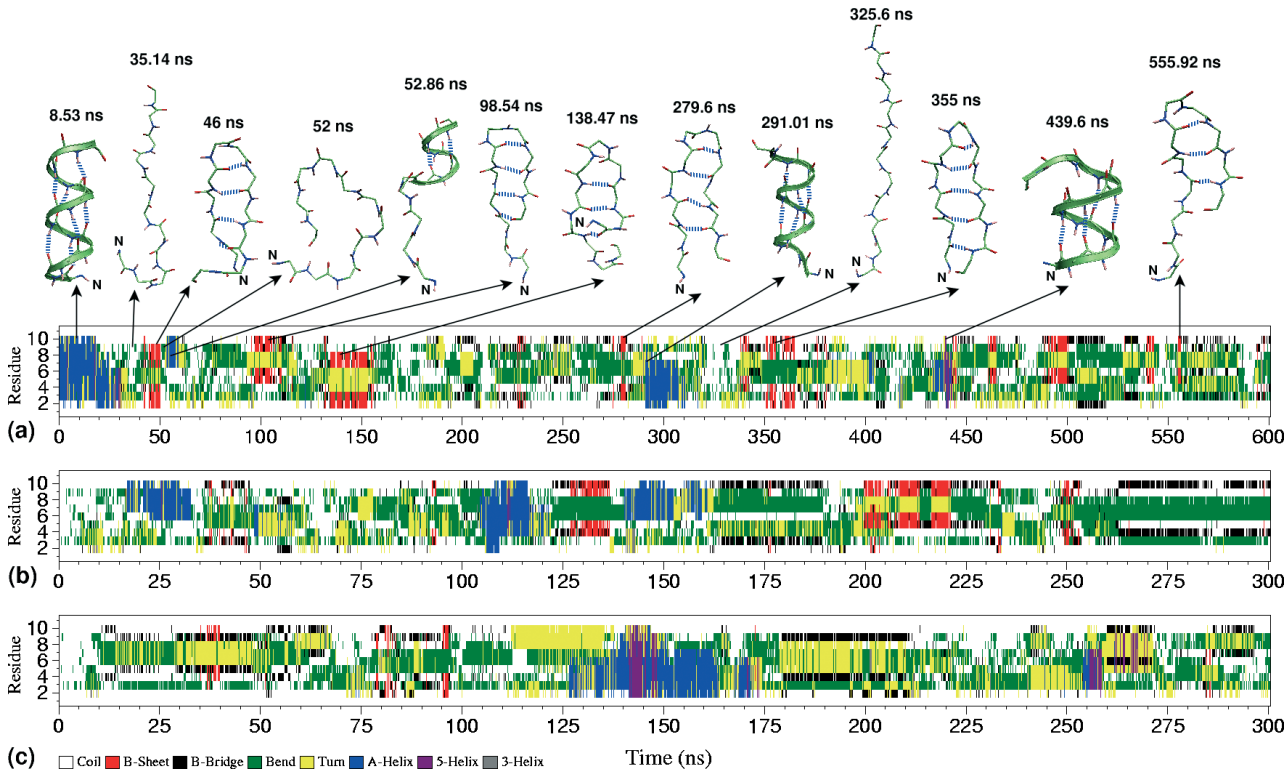

Figure 3 Variation of secondary structure as a function of time in AH simulation (panel a), PP simulation (panel b) and ET simulation (panel c). On the *x*-axis, simulation time is represented in nanoseconds. On the *y*-axis, secondary structure of the residues with appropriate color code is shown. The blue color represents helical conformation, whereas yellow and green colors represent turn and bend conformations, respectively. Sheet conformation is indicated by red color. In this plot, β-hairpins are indicated by a vertical band of red-yellow-red colors. β-Bridges are indicated by a pair of black bands. Conformations sampled at select times are shown above panel-a in AH simulation. The *N*-terminus of the conformations is indicated by the letter '*N*'. These conformations highlight the fact that the peptide samples both the α-helical and β-hairpin conformations along with random coil and bends and turns. In the legend, the 3-helix and 5-helix indicated, represent 3_{10}-helix and π-helix, respectively.

and turns conformations are observed. This family of bends and turns conformations developed into a hairpin conformation (HR1) at about 45 ns (Figures 3 and 4). Shortly thereafter, the hairpin conformation is lost and the bends and turns conformations appear again, which develop into a *C*-terminal partial helix at around 53 ns. Similar pattern of events are observed in all the simulations and throughout the length of a given simulation. A helix or hairpin always unfolds into a family of turns and bends conformations. Further, folding of the peptide into a helix or hairpin is always preceded by bends and turns conformations. Thus, the turns and bends conformations seem to be the intermediate conformational forms in the formation of a helix or hairpin.

The major secondary structural elements observed are turns, bends, helices and hairpins (Figure 3). Sometimes, a partial 3_{10}-helix and a π-helix (partial or full) are observed in the simulations. It may be noted that the coil conformations without any hydrogen bonds are also observed. In all the simulations, there is a reversible formation of a helix and hairpin. The helices and hairpins are separated in time by a family of bends and turns conformational forms. There are a total of eight different types of hairpins observed over the three simulations reported. They are classified into various hairpins, depending upon the residues in the turn region and the interstand hydrogen bonding patterns and are named from HR1 to HR8 (Figure 4). Full and partial α-helical conformations are observed in the simulations. In a full helix, all the residues except those at the termini, sample α-helical conformation. In partial helices, four to five consecutive residues sample α-helical conformation. The partial helix can appear at the *N*- or *C*-terminal end or in the middle region of the peptide.

Formation of β-Hairpin is Mediated by Turns and Bends Conformations

In all the simulations, various types of hairpins are observed as indicated earlier. They differ in terms of the residues in the turn region and the interstrand hydrogen bonding patterns. In AH simulation, six different hairpins are observed (HR1, HR2, HR3, HR4, HR5 and HR6); in PP simulation also, five different hairpins are observed (HR2, HR4, HR6, HR7 and HR8) and in ET simulation, two different types of hairpins

DOI: 10.1002/psc

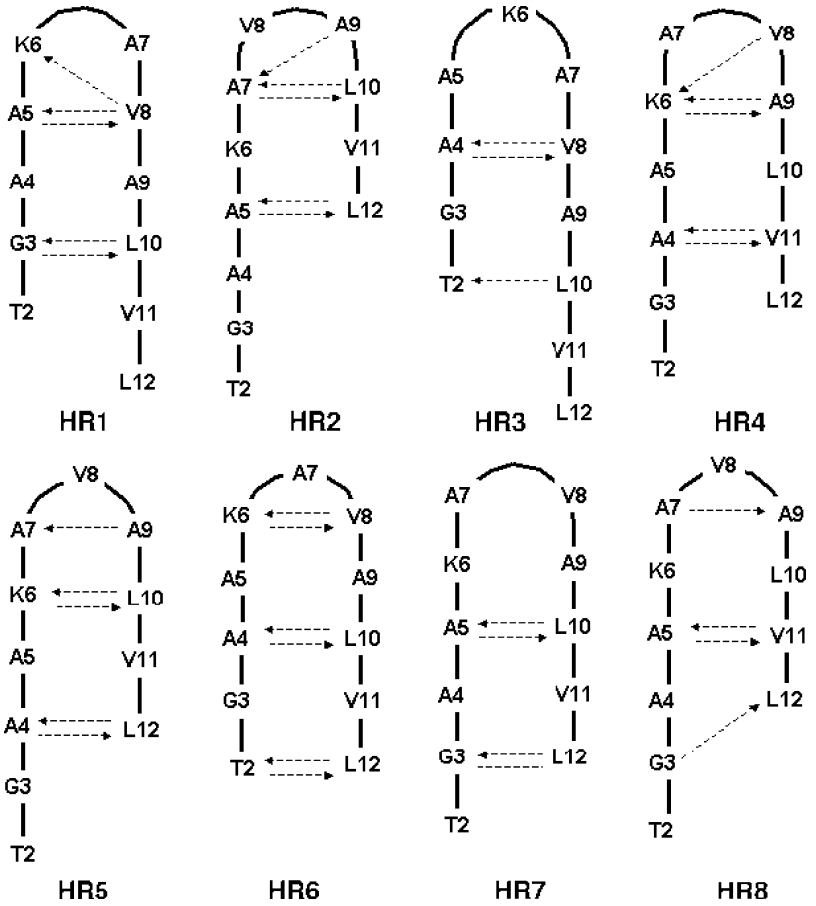

Figure 4 Schematic representation of various hairpins observed in all the simulations. The arrows indicate hydrogen-bonding patterns. The tail of the arrow is at the backbone NH, whereas the head of the arrow is at CO group.

are observed (HR5 and HR6) (Figure 4). We note that the residues, A7 and V8, participate more often in the turn region of the hairpin.

The folding of the hairpin is initiated either from a compact structure or from an unfolded random coil state. An example of the folding of hairpin HR1 in AH simulation from a relatively unfolded conformational state is shown in Figure 5. This hairpin conformation persists for about 5.4 ns. Subsequently, the hairpin conformation is lost and bends and turns form. Similar events are observed when a hairpin conformation forms in all the simulations.

For most of the hairpins, the common mechanism of hairpin formation can be explained in the following steps. (i) Formation of the compact state held by hydrophobic interactions (manifested as bends and turns conformation in DSSP plot in Figure 3). (ii) In the compact conformational state, reorganization takes place to form a new conformation that is closer to a hairpin like conformation, where the peptide takes a U-shaped conformation with its ends closer and the middle region of the U-shaped conformation is held by a hydrophobic cluster formed by two to three hydrophobic residues. That is, the relative alignment of strands takes place first without the development of interstrand hydrogen bonds. (iii) Following this, hydrogen bonds begin to form either from the tail region or from the middle region. When two backbone hydrogen bonds form from the opposite strands in a loop like compact structure, it appears as a bridge conformation in DSSP plot [33]. In a few cases, the formation of a turn takes place prior to the appearance of tail or middle hydrogen bonds. In most of the cases, tail hydrogen bonds appear first, followed by middle hydrogen bonds and then the turn hydrogen bonds. The advantage of this mechanism is that, subsequent to the formation of a compact state, the reduction in entropy is less for the formation of a hairpin.

Interconversion Between Hairpin to α-Helix is also Mediated by Turns and Bends Conformations

Helix formation is also initiated from a compact state consisting of bends and turns. It is always preceded and succeeded by turns or short 3_{10}-helical forms. In all the simulations, helix formation is initiated either from the *N*- or *C*-terminal ends or from the middle region of the peptide leading to a partial helix or full helix.

An example of helix folding mechanism observed in PP simulation is shown in the Figure 6. Here, the

DOI: 10.1002/psc

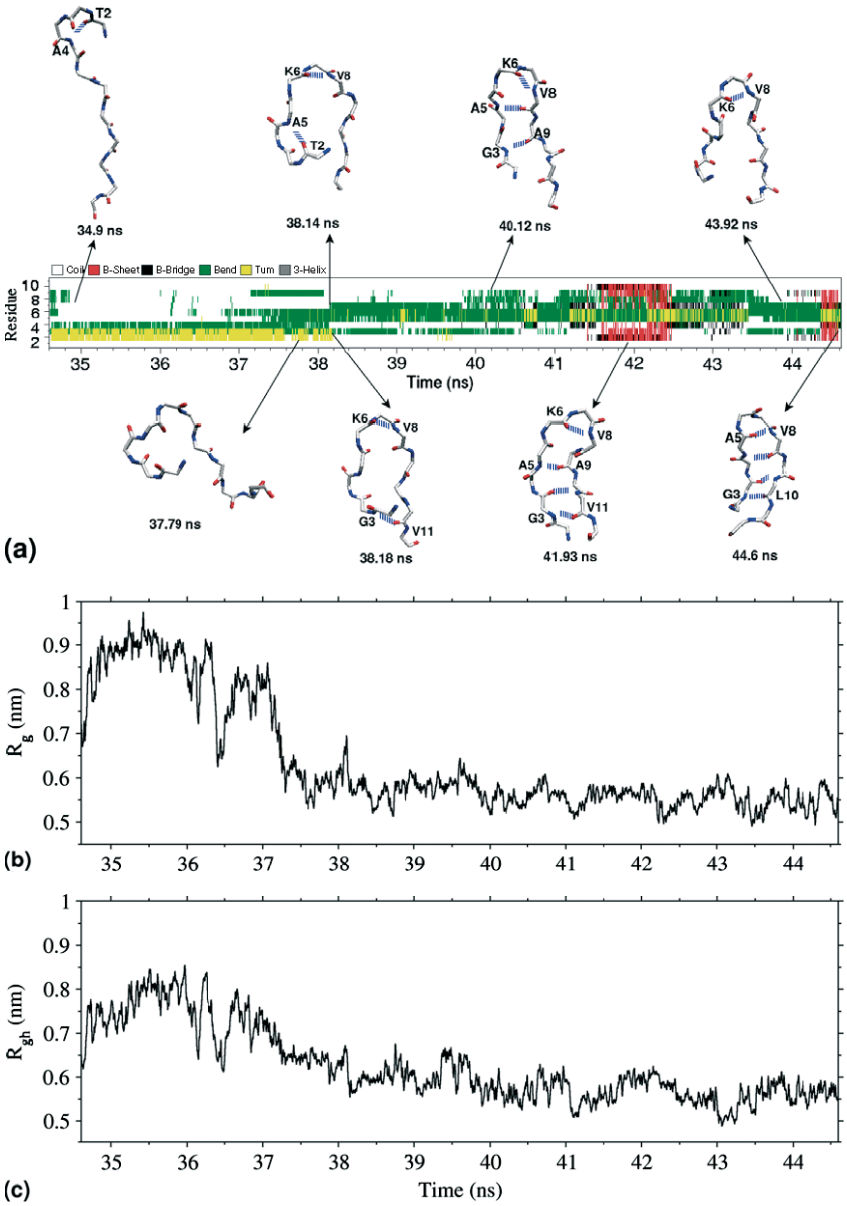

Figure 5 Folding of HR1 hairpin in AH simulation starting from an essentially unfolded state. Middle panel of (a) shows the variation of secondary structures as a function of time in DSSP plot. The top and bottom rows show the backbone conformations at select times. As can be seen from snap shots, during folding, formation of turn structures involving K6, A7 residues as corner residues plays an important role in the folding of β-hairpin HR1 starting from essentially unfolded state. Folding of β-hairpin is accompanied by a decrease in R_g (panel b) and R_{gh} (panel c) values. For visual clarity, running averages over 20 data points (corresponding to 10 ps) are plotted in panels (b) and (c).

peptide starts with a hairpin conformation and ends in an α-helix conformation. Similar events are observed for the formation of a helix from a random coil conformation or from a family of bends and turns conformations. Though it is not shown explicitly in Figure 6(a), formation of short 3_{10}-helix precedes the formation of a helix in most of the cases. Occasionally the regular α-helix interconverts into π-helix, which is sampled for a shorter time of about $\sim$3 ns. Frequently, interconversion between (i + 3, i) and (i + 4, i) hydrogen bonds is observed as steps in the growth of the partial helix. Such interconversion is also observed during helix initiation. It is worth noting that during such interconversions, bifurcated hydrogen bonds are observed most of the time, i.e. (i + 3, i) and (i + 4, i) hydrogen bonds are observed simultaneously. From the Figure 6(a) (middle panel), it is also clear that helix formation is preceded by a sampling of bends and turns conformations, which are essentially compact structures. This can be seen from the plots of R_g and R_{gh} in Figure 6(b) and Figure 6(c). It is worth noting that during the entire transition from α-helix to β-hairpin, R_{gh} essentially remains constant at about 0.57 nm. For comparison, the random coil state has R_g and R_{gh}

DOI: 10.1002/psc

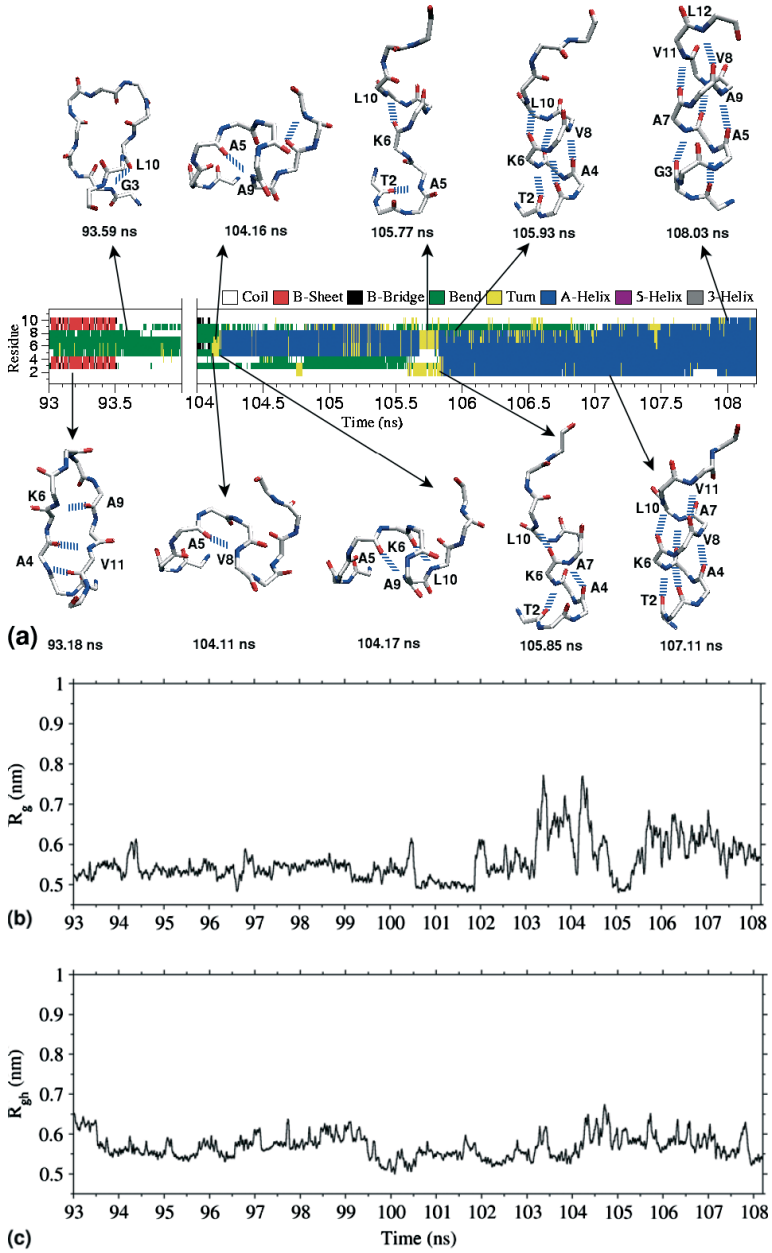

Figure 6 The transition from a β-hairpin to α-helix in PP simulation. In the middle panel of (a), the variation of secondary structures as a function of time is shown. In the top and bottom rows, conformations at select times are shown in backbone representation. Note that in the middle panel, there is a gap between 94 and 104 ns, which essentially represents bends and turns conformations. The mechanism involves unfolding of β-hairpin conformation into a family of bends and turns conformations. As can be seen, both (i + 3, i) and (i + 4, i) turns play an important role in the formation of helix. Note that during the entire transition, the R_{gh} essentially remains constant (0.57 ± 0.03 nm), as can be seen in panel (c). The backbone R_g (panel b), also has not varied much during the transition except at about 103.2 to 105 ns, where it sampled to a relatively higher R_g value owing to partial unfolded forms involving residues TGA at *N*-terminus. However, around the same time, hydrophobic side chains maintained a similar compactness and hence R_{gh} has not varied much. For visual clarity, running averages over 20 data points (corresponding to 10 ps) are plotted in panels (b) and (c).

values of 1.03 ± 0.06 and 0.76 ± 0.04 nm, respectively. This suggests that the intermediate compact bends and turns conformational forms, α-helix and β-hairpin have similar hydrophobic interactions. The intermediate compact states may, therefore, be considered as a collapsed states governed largely by hydrophobic interactions.

Turn and Bend Potential is Spread Throughout the Sequence

In all the simulations, turns and bends conformations are observed most of the time, as is evident from the DSSP plot (Figure 3). From Figure 7, it is clear that the turn and bend propensity is spread throughout

DOI: 10.1002/psc

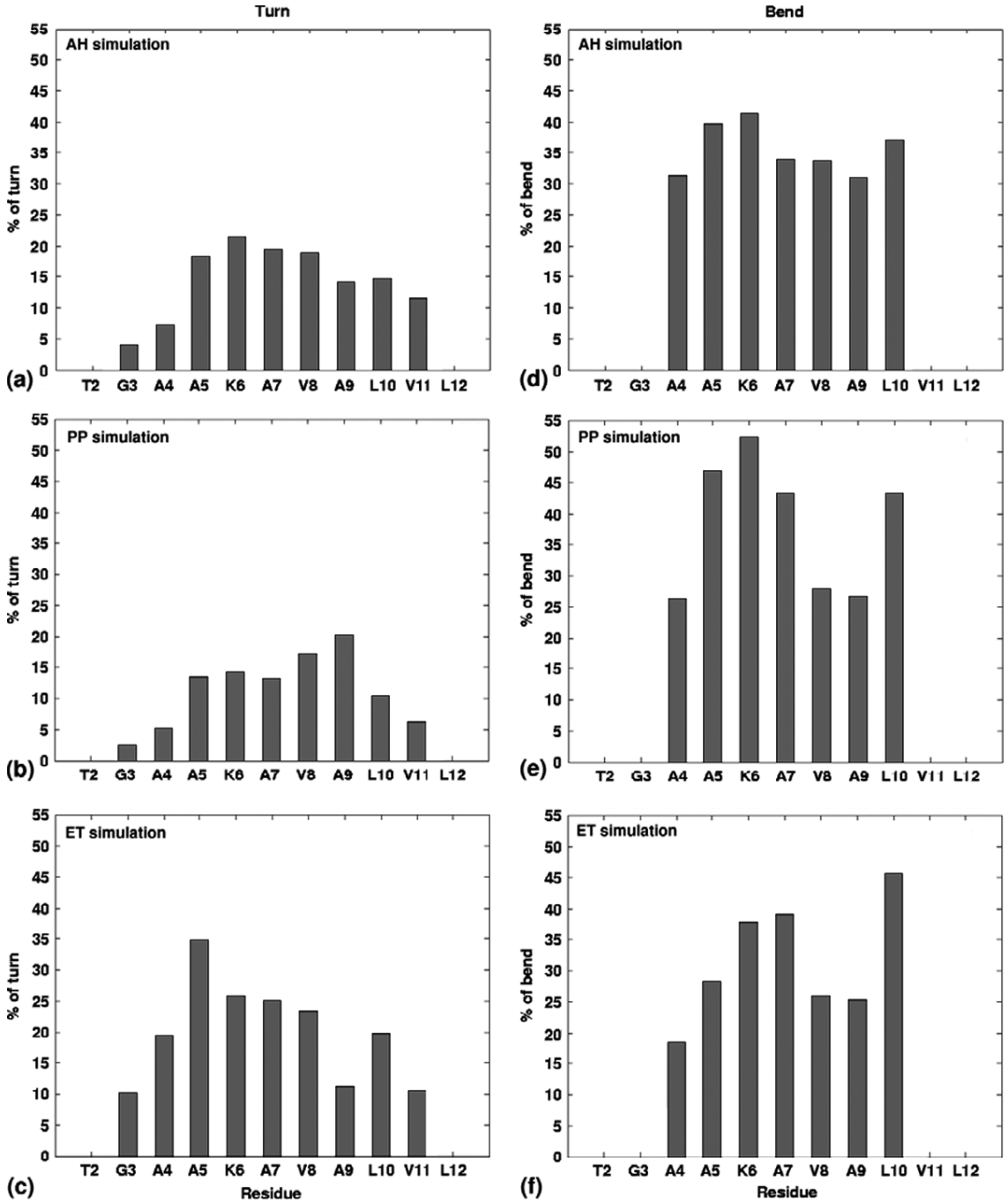

Figure 7 The observed propensity for turn and bend conformations in AH simulation (a) and (d), in PP simulation (b) and (e) and ET simulation (c) and (f). The residues in the peptide are shown on *x*-axis and the participation of a given residue in a turn or bend conformation as a percentage of the simulation time is indicated along the *y*-axis. Bend and turn assignments are taken from DSSP plots (Figure 3). It can be seen that bends have a relatively higher propensity than turns.

the sequence. Note that a turn can be either a two-residue (i + 3, i) hydrogen bonded turn or a three-residue (i + 4, i) hydrogen bonded turn.

Bends and Turns Frequently Form from Coil State

A bend conformation involves five residues and the middle residue is said to adopt a bend conformation [33] (See Section on Methods). Such bend conformations are observed frequently in the simulations. Subsequent to the formation of a bend, a turn with i, i − 1 or i, i + 1 corner residues can form. A bend conformation can appear anywhere in the sequence except at the *N*-terminal two residues and *C*-terminal two residues in view of its definition [33].

In all the simulations, bend conformation is the first one to form, starting from completely unfolded state, where all the residues of the peptide sample

DOI: 10.1002/psc

coil conformation. Bend conformation is initiated at one residue, which then propagates to include more residues or is simultaneously adopted by more than one residue to form a set of bends. In a bend conformation, the side chains of the hydrophobic residues could come in contact and form a hydrophobic cluster. This gives rise to a compact state. The hydrophobic interactions seem to stabilize the compact state. There is no specific conformation for the compact structure; it can consist of several bends or turns conformations. A set of bends and turns give rise to a compact state/collapsed state that serves as starting point for the initiation of both helix and hairpin conformations (Figure 6).

The Distribution of Conformational Energies is Gaussian with Helix Sampling Lowest Energy Followed by Hairpins and Coils

The potential energy samples a Gaussian distribution in all the simulations. The full helix conformation is sampled by the lower energy end of the histogram with an average energy of 1250 kJ/mol, whereas the coil conformation occupies the higher energy end of the histogram with an average energy of 1425 kJ/mol. The middle peaks of the histogram are populated by hairpins, partial helices and other compact states formed by bridges, bends and turns. Thus, α-helix and β-hairpin may be considered as fluctuations in a collapsed state.

DISCUSSION

Bends and Turns Conformations Sampled More Often

The bends and turns conformations are sampled more often than the helix and hairpin conformations in the simulations. They have intermediate potential energies between the lower energy of full helix and the higher energy of coil conformation as discussed earlier. Though the full helix has a lower energy, it is not sampled frequently owing to high cost of entropy. In contrast, bends/turns, bridges and so on have relatively higher energy, yet they are sampled more frequently owing to entropic advantage: there are many ways in which bends/turns can form.

On why Different Hairpins are Observed

In our simulations, various types of hairpins are observed (Figure 4). These hairpins differ in the residues of the turn region and in the interstrand hydrogen bonding patterns. As discussed earlier in the Results, the turn/bend potential is spread throughout the sequence (Figure 7); as a consequence, a turn can form anywhere in the middle region of the sequence that would serve as a turn in the hairpin structure. This leads to various hairpins as observed. Further, the turn propensity can also lead to a β-turn, which gives rise to an α-turn and finally to helical structures.

General Mechanism for Conformational Transitions Observed in the Simulations

As discussed earlier, the transition between α-helix and β-hairpin or vice versa is mediated through a family of bends and turns conformations, which can be considered as a collapsed state as discussed in Results, earlier. For example, an α-helix unfolds into bends and turns conformations, which further develop into a β-hairpin conformation. Similarly, a β-hairpin unfolds into a family of bends and turns conformations before adopting an α-helical conformation. Thus, it appears necessary for the sequence to adopt bends and turns conformations before proceeding to an α-helix conformation or to a β-hairpin conformation. This is summarized schematically in Figure 8. Sometimes, turns and bends conformations unfold into a coil conformation, which again regains bends and turns conformations. A coil conformation can become a hairpin conformation or helix conformation only through bends and turns conformations. Thus, conformations with bends and turns play an important role as an intermediate in the conformational transitions.

Other Simulations and Experimental Studies also Highlight the Importance of Turns (and Bends) Conformations in the Generation of α-Helical Conformation

From experimental ^{1}H NMR studies of the peptide fragment TGAAKA, which is a part of our current sequence TGAAKAVALVL, it was concluded that the residues of the peptide sample α-helical and neck region of the ϕ, ψ map with reduced conformational entropy [35]. It is further concluded that the peptide has a potential for turn conformational forms. These experimental observations correlate well with our observation that bends and turns propensity is spread throughout the sequence. Further, there is considerable sampling of α-helical and neck regions of the Ramachandran ϕ, ψ map for most of the residues of the peptide (data not shown). Sampling in these regions is known to lead to turn conformational forms.

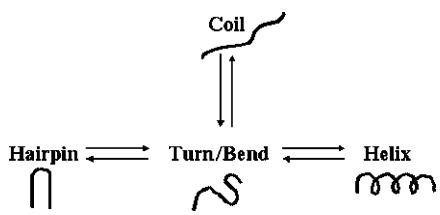

Figure 8 Schematic diagram of a general mechanism for the formation of helix and hairpin conformations.

DOI: 10.1002/psc

For example, these regions in the ϕ, ψ map are sites for type-I and type-III turn conformational forms.

In a ^{1}H NMR experimental study of a sequence from myohemerythrin, it was found that the peptide samples an ensemble of turn conformations, which readily goes into an α-helix in the presence of trifluroethanol [36]. The ensemble of turn conformations is termed as a nascent helix to indicate its potential to adopt helical conformation. In our simulations, we find a verification of this concept that a family of turns (and bends) lead to helical conformation.

Wu and Wang [37] in their study on a 16-residue alanine-based polypeptide in explicit water observed that turns and 3_{10}-helices play an essential role in the folding of an α-helix. A role for 3_{10}-helix and β-turn in helix initiation is observed by Monticelli *et al.* [38]. In our simulation also, initiation of the helix is mediated by turns and 3_{10}-helical fragments. Pal *et al.* [39] observed that the initiation of an α-helix or a 3_{10}-helix involve the formation of an isolated β-turn. In our simulations, we find that bends and turns conformations lead to helical and hairpin conformations; the formation of a helix is more often preceded by a 3_{10}-helical conformation or β-turns. Sung and Wu [40] find that the (i + 3, i) hydrogen bond is frequently observed during helix folding and unfolding. In a study by Sung [41] on helix folding simulation from various initial conformations, polyalanine folds into an α-helical conformation where β-bends are observed as intermediates during the helix nucleation. In that study, it was also observed that the helix hydrogen bond interconverts between the (i + 4, i) and the (i + 3, i) type, and the (i + 3, i) hydrogen bonds occurred more frequently during helix propagation. Sundaralingam and Sekharudu [17] in their study of hydrated helical segments from the crystal structures of the protein find that a variety of turn conformations are implicated in the folding and unfolding of α-helices. In our simulations also, similar observations are made.

Daidone *et al.* [9] and Simona *et al.* [10] in their MD simulations study on fibrillogenic peptides observed that the transition from α-helical to β-structure requires the peptide to sample intermediate β-bend structures. Similar intermediate β-bend structures are also observed in our simulations during the transition from α-helix to β-hairpin and vice versa.

Comparison to Other Simulations on the Mechanism of Hairpin Formation

In the study of β-hairpin folding, Sung [42] finds that the peptide first folds into a compact U-shaped conformation, stabilized by hydrophobic interactions. Following a reorganization of this compact state, β-hairpin hydrogen bonds form. In our simulations also, most often the peptide adopts a U-shaped conformation before the development of hairpin conformation (Figure 5) and hydrophobic interactions seem to be playing a role as observed by Sung [42].

Shorter Peptide Fragments from the Sequence TGAAKAVALVL Display Structural Ambivalence in Known Structures of Proteins which are Dissimilar to Glyceraldehyde-3-Phosphate Dehydrogenase

Table 1 gives peptide fragments from the sequence TGAAKAVALVL and their occurrence in various protein structures, which are dissimilar to glyceraldehyde-3-phosphate dehydrogenase along with the conformations sampled by them. From the Table 1, it can be seen, for example, the fragment AKAV samples β-strand in the protein with PDB code 1UKK and α-helix in protein with PDB code 1JDS. As another example, the fragment VALVL samples helical conformation in protein with PDB code 1JN0, whereas the same sequence samples β-strand conformation in protein with PDB code 1UUS. This shows the conformational plasticity of the sequence to adopt helical and strand conformation in crystal structures. Table 1 gives many such examples for overlapping peptide fragments from the sequence.

The Peptide Displays Structural Ambivalence

In the crystal structure of the protein, glyceraldehyde-3-phosphate dehydrogenase, the peptide under study adopts helical conformation. However, in our simulation of the peptide in explicit SPC water, the peptide samples α-helical as well as β-hairpin conformations as discussed in the Results. Recently, an algorithm has been developed to predict the propensity of a given sequence to be amyloidgenic based upon hidden propensity to sample β-strand conformation. These hidden propensities are computed at low and high tertiary contact (TC) level [43]. Our peptide sequence is predicted by the algorithm [43] to take up α-helical conformation at low TCs and β-structure at high TCs. Thus, the peptide is predicted to be 'structurally ambivalent' and this property is observed in our simulations.

The sequence has plenty of Alanines, a couple of Valines and Leucines, which play a role in determining its conformational plasticity. It is known that Valine, Isoleucine, Leucine and Alanine are prevalent among the chameleon peptides [4,5]. In most of the SAPs, one of the two strong helix former (A or E) and one of the two strong strand formers (V or I) appear simultaneously [5]. It is worth noting that Leucine (L) is known to be both helix and strand former with a strong propensity. Zhou *et al.* [5] also find that dipeptides consisting of a strong helix former and a strong strand former (LL, IA, AV, LV, LI, AI, EL and EI) occur more frequently in SAPs. The *C*-terminal end of our peptide (AVALVL) has many such dipeptide combinations.

On why the Sequence TGAAKAVALVL Displays Structural Ambiguity

The peptide has many hydrophobic residues. We can, therefore, expect its conformation to be determined

DOI: 10.1002/psc

largely by hydrophobic interactions. The R_{gh} values of helix and hairpin are similar (≈0.55 nm), indicating that in both the conformations the hydrophobic residues have similar compact distribution. The Lennard-Jones energies of helix and hairpin are similar (~178 kJ/mol). The Lennard-Jones energies of hydrophobic residues alone for helix and hairpin conformations are within 19 kJ/mol of each other (corresponds to energy of a hydrogen bond). If we take Lennard-Jones (hydrophobic) energies as a measure of how closely hydrophobic residues are associated, we may again infer that hydrophobic residues have a similar compact distribution in both helix and hairpin conformations. We, therefore, suggest that the conformational features of TGAAKAVALVL are determined largely by hydrophobic interactions and hydrophobic interactions are similar in helix and hairpin conformations (Figure 6). Hence, these conformations are sampled with similar frequency. The similarity in hydrophobic interactions in helix, hairpin and other compact conformations may be related to the distribution of hydrophobic residues in the sequence.

The Sequence TGAAKAVALVL Adopts Helical Conformation in Glyceraldehyde-3-Phosphate Dehydrogenase due to its Low Conformational Energy

In a simulation study on polyalanine, it is observed that a metastable β-hairpin intermediate occurs and the potential energy of β-hairpin is higher than the α-helix [44]. In our case also, it is the α-helix that has a lower potential energy. This also perhaps explains why the sequence adopts helical conformation in glyceraldehyde-3-phosphate dehydrogenase.

Implication for Protein Folding

In glyceraldehyde-3-phosphate dehydrogenase, the sequence TGAAKAVALVL adopts a helical conformation. However, as we have seen in the MD simulation, the peptide samples both α-helical and β-hairpin conformations. This implies that a selection has to be made during protein folding for helical conformation. Those conformers with β-hairpin need to convert to helical conformation so as to proceed with the folding.

The hydrophobic collapse model of the protein folding [45–47] states that a nonspecific collapse takes place during early folding and within the collapsed state, secondary and tertiary structures develop. What is observed in our simulation is that a compact conformational state consisting of bends and turns forms, starting from unfolded conformational states and within the compact conformational state, the secondary structures α-helix and β-hairpin generate. This observation is analogous to hydrophobic collapse model of protein folding [45–47].

CONCLUSIONS

The peptide TGAAKAVALVL sequence from glyceraldehyde-3-phosphate dehydrogenase displays a structural ambiguity and interconverts between α-helical and β-hairpin conformations in MD simulations over a total simulation time of 1.2 μs, starting from various initial conformations. The conformational transitions are mediated by turns and bends conformations, which are also compact structures. The structural ambiguity displayed by the sequence may be attributed to the distribution of hydrophobic residues in the sequence and similarity of hydrophobic interactions in both helical and hairpin conformations. The residues in the peptide sequence have propensity for bends and turns conformations and many residues from the sequence are also prevalent in chameleon sequences. This also seems to be the cause for the observed structural ambiguity. The full helix conformation has a lower energy than other observed conformations, which is probably why the peptide takes up helical conformation in the crystal structure of the protein glyceraldehyde-3-phosphate dehydrogenase.

Acknowledgements

We thank BRNS for funding the research through a research grant, sanction number 2001/37/18/BRNS/796, sanctioned to YUS and SP thanks BRNS for JRF fellowship.

REFERENCES

1. Wilson IA, Haft DH, Getzoff ED, Tainer JA, Lerner RA, Brenner S. Identical short peptide sequences in unrelated proteins can have different conformations: a testing ground for theories of immune recognition. *Proc. Natl. Acad. Sci. U.S.A.* 1985; **82**: 5255–5259.
2. Cohen BI, Presnell SR, Cohen FE. Origins of structural diversity within sequentially identical hexapeptides. *Protein Sci.* 1993; **2**: 2134–2145.
3. Kabsch W, Sander C. On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. *Proc. Natl. Acad. Sci. U.S.A.* 1984; **81**: 1075–1078.
4. Mezei M. Chameleon sequences in the PDB. *Protein Eng.* 1998; **11**: 411–414.
5. Zhou X, Alber F, Folkers G, Gonnet GH, Chelvanayagam G. An analysis of the helix-to-strand transition between peptides with identical sequence. *Proteins* 2000; **41**: 248–256.
6. Kuznetsov IB, Rackovsky S. On the properties and sequence context of structurally ambivalent fragments in proteins. *Protein Sci.* 2003; **12**: 2420–2433.
7. Minor DL Jr, Kim PS. Context-dependent secondary structure formation of a designed protein sequence. *Nature* 1996; **380**: 730–734.
8. Young M, Kirshenbaum K, Dill KA, Highsmith S. Predicting conformational switches in proteins. *Protein Sci.* 1999; **8**: 1752–1764.
9. Daidone I, Simona F, Roccatano D, Broglia RA, Tiana G, Colombo G, Di Nola A. Beta-hairpin conformation of fibrillogenic peptides: structure and alpha-beta transition mechanism revealed by molecular dynamics simulations. *Proteins* 2004; **57**: 198–204.

10. Simona F, Tiana G, Broglia RA, Colombo G. Modeling the alpha-helix to beta-hairpin transition mechanism and the formation of oligomeric aggregates of the fibrillogenic peptide Abeta(12–28): insights from all-atom molecular dynamics simulations. *J. Mol. Graph. Model.* 2004; **23**: 263–273.
11. Dima RI, Thirumalai D. Exploring the propensities of helices in PrP(C) to form beta sheet using NMR structures and sequence alignments. *Biophys. J.* 2002; **83**: 1268–1280.
12. Gross M. Proteins that convert from alpha helix to beta sheet: implications for folding and disease. *Curr. Protein Pept. Sci.* 2000; **1**: 339–347.
13. Hamada D, Kuroda Y, Tanaka T, Goto Y. High helical propensity of the peptide fragments derived from beta-lactoglobulin, a predominantly beta-sheet protein. *J. Mol. Biol.* 1995; **254**: 737–746.
14. Hamada D, Segawa S, Goto Y. Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein. *Nat. Struct. Biol.* 1996; **3**: 868–873.
15. Kuroda Y, Hamada D, Tanaka T, Goto Y. High helicity of peptide fragments corresponding to beta-strand regions of beta-lactoglobulin observed by 2D-NMR spectroscopy. *Fold. Des.* 1996; **1**: 255–263.
16. Dobson CM, Evans PA, Radford SE. Understanding how proteins fold: the lysozyme story so far. *Trends Biochem. Sci.* 1994; **19**: 31–37.
17. Sundaralingam M, Sekharudu YC. Water-inserted alpha-helical segments implicate reverse turns as folding intermediates. *Science* 1989; **244**: 1333–1337.
18. Mu Y, Kosov D, Stock G. Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments. *J. Phys. Chem. B* 2003; **107**: 5064–5073.
19. Woutersen S, Hamm P. Isotope-edited two-dimensional vibrational spectroscopy of a short alpha-helix in water. *J. Chem. Phys.* 2001; **114**: 2727–2737.
20. Eker F, Cao X, Nafie L, Schweitzer-Stenner R. Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR, and VCD spectroscopy study. *J. Am. Chem. Soc.* 2002; **124**: 14330–14341.
21. Shi Z, Olson CA, Rose GD, Baldwin RL, Kallenbach NR. Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl. Acad. Sci. U.S.A.* 2002; **99**: 9190–9195.
22. Kentsis A, Mezei M, Gindin T, Osman R. Unfolded state of polyalanine is a segmented polyproline II helix. *Proteins* 2004; **55**: 493–501.
23. Mezei M, Fleming PJ, Srinivasan R, Rose GD. Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins* 2004; **55**: 502–507.
24. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 1997; **18**: 2714–2723.
25. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In *Inter Molecular Forces*, Pullman B (ed.). Reidel Publishing Company Dordrecht: The Netherlands, 1981; 331–342.
26. Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 1993; **98**: 10089–10092.
27. Essmann U, Perera L, Berkowitz M, Darden T, Lee H, Pedersen L. A smooth particle mesh Ewald method. *J. Chem. Phys.* 1995; **103**: 8577–8593.
28. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 1997; **18**: 1463–1472.
29. Berendsen H, Postma J, DiNola A, Haak J. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984; **81**: 3684–3690.
30. Lindahl E, Hess B, van der Spoel D. Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* 2001; **7**: 306–317.
31. van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. vdf Hochschulverlag AG an der ETH: Zürich, Groningen, 1996.
32. Scott WRP, Huenenberger P, Tironi I, Mark A, Billeter S, Fennen J, Torda A, Huber T, Krueger P, van Gunsteren W. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* 1999; **103**: 3596–3607.
33. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983; **22**: 2577–2637.
34. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J. Mol. Graphics* 1996; **14**: 27–28, 33–38.
35. Sasidhar YU, Ramakrishna V. Conformational features of a hexapeptide model Ac-TGAAKA-NH2 corresponding to a hydrated alpha helical segment from glyceraldehyde 3-phosphate dehydrogenase: implications for the role of turns in helix folding. *Indian J. Biochem. Biophys.* 2000; **37**: 34–44.
36. Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin. *J. Mol. Biol.* 1992; **226**: 795–817.
37. Wu X, Wang S. Helix folding of an alanine-based peptide in explicit water. *J. Phys. Chem. B* 2001; **105**: 2227–2235.
38. Monticelli L, Tieleman DP, Colombo G. Mechanism of helix nucleation and propagation: Microscopic view from microsecond time scale MD simulations. *J. Phys. Chem. B* 2005; **109**: 20064.
39. Pal L, Chakrabarti P, Basu G. Sequence and structure patterns in proteins from an analysis of the shortest helices: implications for helix nucleation. *J. Mol. Biol.* 2003; **326**: 273–291.
40. Sung SS, Wu XW. Molecular dynamics simulations of synthetic peptide folding. *Proteins* 1996; **25**: 202–214.
41. Sung SS. Helix folding simulations with various initial conformations. *Biophys. J.* 1994; **66**: 1796–1803.
42. Sung SS. Monte Carlo simulations of beta-hairpin folding at constant temperature. *Biophys. J.* 1999; **76**: 164–175.
43. Yoon S, Welsh WJ. Detecting hidden sequence propensity for amyloid fibril formation. *Protein Sci.* 2004; **13**: 2149–2160.
44. Ding F, Borreguero JM, Buldyrey SV, Stanley HE, Dokholyan NV. Mechanism for the alpha-helix to beta-hairpin transition. *Proteins* 2003; **53**: 220–228.
45. Dill KA. Theory for the folding and stability of globular proteins. *Biochemistry* 1985; **24**: 1501–1509.
46. Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. Principles of protein folding–a perspective from simple exact models. *Protein Sci.* 1995; **4**: 561–602.
47. Dill KA. Dominant forces in protein folding. *Biochemistry* 1990; **29**: 7133–7155.